

TUTKIMUSAINEISTOJEN HALLINTA JA ETIIKKA

FT, tiiminvetäjä Katja Fält

Tampereen yliopiston kirjasto, Avoimen tieteen palvelut

katja.falt@tuni.fi

ORCID ID: <https://orcid.org/0000-0002-6172-5377>

10.11.2020



This work is licenced under
Creative Commons Attribution
4.0 International License.

TUTKIMUSDATA/-AINEISTO

Tutkimuksen tuottamaa tai sen tarpeisiin tuotettua tai kerättyä faktuaalista tietoa (numeerinen, tekstuaalinen, visuaalinen, auditiivinen)

Mitä tutkimusdata on, määrittyy käytön perusteella ja lähes mikä tahansa digitaalinen objekti voi hahmottua tutkimusdatana, kun sitä käytetään tutkimuskysymykseen vastaamisessa

DATA – TUTKIMUSTULOSTEN PERUSTA



“Research data are the evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical).”

Concordat on Open Research Data, published on 28 July 2016

<https://www.ukri.org/files/legacy/documents/concordatonopenresearchdata-pdf/>

MITÄ ON AINEISTONHALLINTA?

**AINEISTONHALLINNALLA
TARCOITETAAN SITÄ,
ETTÄ**

- aineistot on luotu, tallennettu, kuvailtu ja järjestetty suunnitelmallisesti
- aineistojen käyttöoikeudet ja elinkaari on suunniteltu ja yhteisesti sovittu

**AINEISTONHALLINNAN
TARCOITUS ON
VARMISTAA, ETTÄ**

- tutkimusaineistojen suhteen noudatetaan hyvää tieteellistä käytäntöä
- tutkimusaineisto ei vaarannu missään vaiheessa
- aineiston jatkokäyttö on mahdollista

AINEISTONHALLINTAA EDELLYTTÄÄ

KOTI- ORGANISAATIO

Tampereen korkeakouluyhteisön Avoimen tieteen linjaus

- Tutkimusdata on lähtökohtaisesti yhteiskäyttöistä ja avointa.
- Huolellisesti koottu, järjestetty, kuvailtu ja uudelleen- ja jatkokäyttöön avattu tutkimusdata luetaan tutkijalle tieteelliseksi ansioksi
- Tutkija laatii tutkimuksestaan jo suunnitteluvaiheessa aineistohallintasuunnitelman

TUTKIMUKSEN RAHOITTAJAT

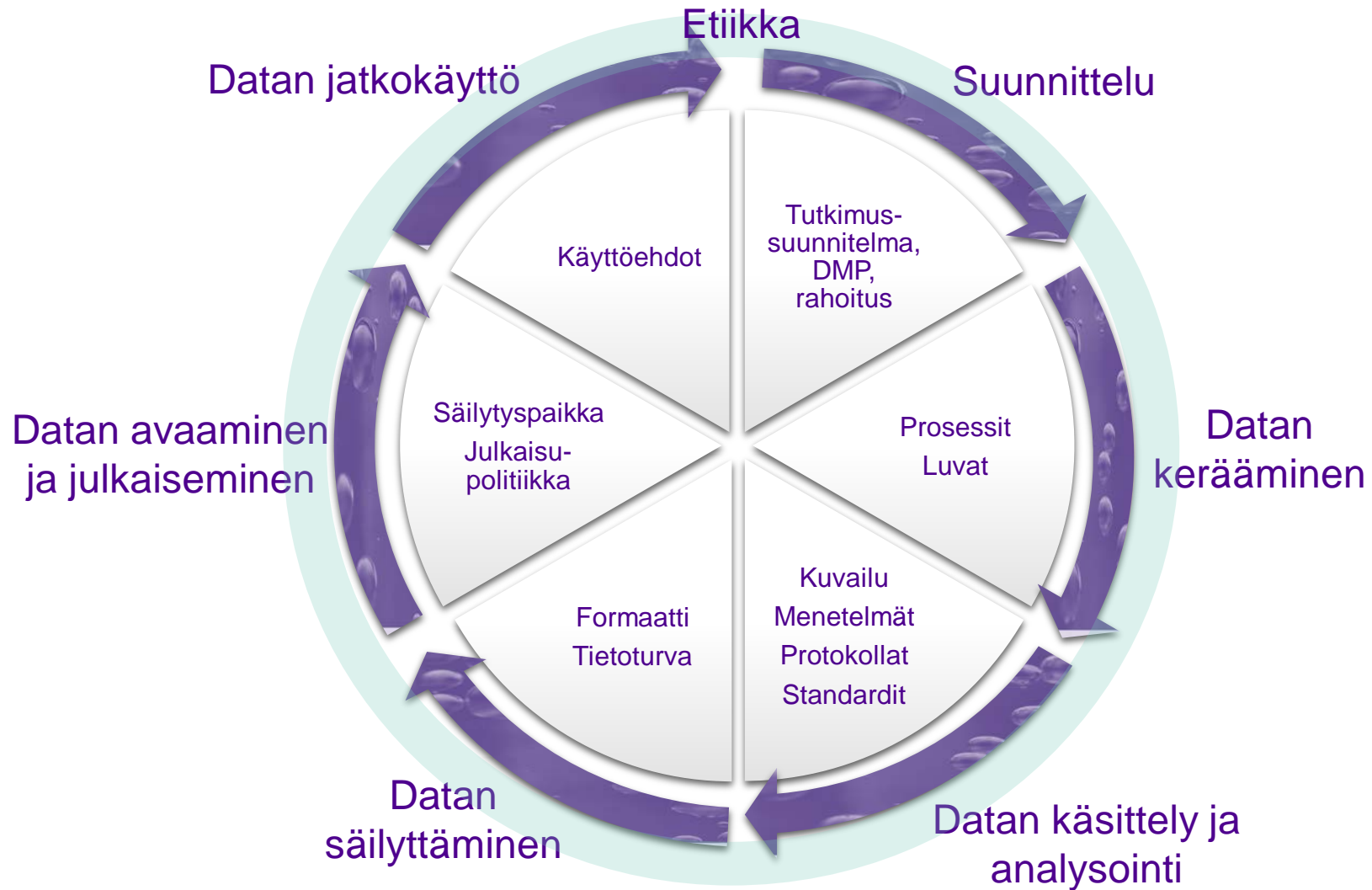
- **Suomen Akatemia** vaatii aineistohallintasuunnitelmaa, data tulee saattaa avoimesti saataville tulosten julkaisemisen jälkeen
- **EU's Horizon2020:** Open access on oletusarvo
- **Business Finland:** Data tulee hyödyntää tulevaisuudessa mahdollisimman tehokkaasti

MIKSI AINEISTONHALLINTAA?

- Tutkimusaineistojen hallinta on osa hyvää tieteellistä käytäntöä
- Voidaan varmistaa vastuullinen ja eettinen tutkimus
- Hyvät aineistohallintatavat ovat oletusarvo ja olennainen osa hyviä tutkijan taitoja!
- Ennaltaehkäisee odottamattomia ongelmia (datan häviäminen/tuhoutuminen, eettiset tai lainsäädännölliset hankaluudet)
- Auttaa hallitsemaan ja järjestämään omaa tutkimusaineistoa
- Aineiston avaaminen, jakaminen ja uudelleenkäyttö on helpompaa
- Voidaan osoittaa, mihin aineistoon tutkimuksen tulokset perustuvat
- Parantaa tutkimuksen läpinäkyvyyttä, luotettavuutta, toistettavuutta ja todennettavuutta
- Aineistohallinnan suunnittelu auttaa seuraamaan rahoittajien linjauksia
- Osoittaa, että ymmärrät tutkimusaineistojesi arvon
- Kertoo osaamisestasi tutkimushankkeen vetäjänä
- Suunnittelussa auttaa aineistohallintasuunnitelma (DMP, data management plan)

Lisätietoa: <https://libguides.tuni.fi/tutkimusaineistojen-hallinta/suunnittele>

AINEISTON ELINKAARI





TUTKIMUSETIIKKA JA AINEISTONHALLINTA - ENNEN TUTKIMUSTA

HYVÄ TIETEELLINEN KÄYTÄNTÖ

- Tutkimukseen sovelletaan tieteellisen tutkimuksen kriteerien mukaisia ja eettisesti kestäviä tiedonhankinta-, tutkimus- ja arviointimenetelmiä
- Tutkimuksessa toteutetaan tieteellisen tiedon luonteeseen kuuluvaa **avoimuutta** ja vastuullista tiedeviestintää tutkimuksen tuloksia julkaistaessa
- Tutkijat ottavat muiden tutkijoiden työn ja saavutukset asianmukaisella tavalla huomioon (viittaaminen)
- Tutkimus suunnitellaan ja toteutetaan ja siitä raportoidaan sekä siinä syntyneet **tietoaineistot tallennetaan** tieteelliselle tiedolle asetettujen vaatimusten edellyttämällä tavalla
- Tarvittavat tutkimusluvut on hankittu ja tietyillä aloilla vaadittava eettinen ennakoarviointi on tehty
- Tutkimushankkeessa tai -ryhmässä sovitaan ennen tutkimuksen aloittamista kaikkien osapuolten oikeudet, tekijyyttä koskevat periaatteet, vastuut ja velvollisuudet sekä **aineistojen säilyttämistä ja käyttöoikeuksia koskevat kysymykset**

Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet (TENK 2019)

Tutkimusaineistojen avoimuus

- Tutkimuksessa kerätyn aineiston tallentaminen muiden tutkijoiden saataville on yksi tapa toteuttaa tieteen avoimuutta
- Avoimuuden aste määritetään aineistokohtaisesti ottaen huomioon sekä tieteen vapaus ja sananvapaus että henkilötietojen ja yksityisyyden suoja
- Tutkimusaineiston avaaminen otetaan huomioon jo tutkimuksen suunnitteluvaiheessa.

Hyvä tieteellinen käytäntö korkeakouluyhteisössä: <https://www.tuni.fi/fi/tutkimus/vastuullinen-tiede/hyva-tieteellinen-kaytanto>

YLEISET EETTISET PERIAATTEET

1. Tutkittavan ihmisarvon ja itsemääräämisoikeuden kunnioittaminen,

2. Aineellisen ja aineettoman kulttuuriperinnön sekä luonnon monimuotoisuuden kunnioittaminen

3. Riskien, haittojen ja vahingoittamisen välttäminen

SELVITÄ HYVISSÄ AJOIN

- Henkilötietojen keräämisen tarve
 - Perehdy korkeakouluyhteisön [tutkimuksen tietosuojapolkuun](#)
 - Valmistele [tutkittavien informointi](#)
 - Laadi tutkimusilmoitus: tutkimuksen aihe ja mistä tutkimuksessa on kyse
 - Laadi tietosuojailmoitus ja riskienarviointi
 - Eettinen osallistumissuostumus tutkittavilta (tietoon perustuva suostumus) ja lupa aineiston jatkokäyttöön
- Tarvitaanko eettinen ennakoarvio?
- Tarvitaanko [tutkimuslupa](#) kohdeorganisaatiolta?
 - TAU: Jos kyseessä on opinnäytetyö tai tutkimus, joka koskee yksittäistä tiedekuntaa, luvan myöntää kyseisen tiedekunnan dekaani
 - Tutkimuslupa myönnetään ensisijaisesti sellaisille tieteellisille ja/tai opinnäytetyöhön sekä kehitystyöhön liittyville aloille, jotka liittyvät yliopiston toimintaan ja kehittämiseen.
- Muiden aineistojen käyttöön liittyvät tekijänoikeudet ja käyttöoikeudet

AINEISTOISTA SOPIMINEN

- Tutkimushankkeessa tai -ryhmässä sovitaan ennen tutkimuksen aloittamista kaikkien osapuolten oikeudet, tekijyyttä koskevat periaatteet, vastuut ja velvollisuudet sekä aineistojen säilyttämistä ja käyttöoikeuksia koskevat kysymykset
- Osoittaa vastuullisuutta ja kunnioitusta tutkittavia kohtaan
- Selventää aineistoihin liittyviä oikeuksia ja auttaa määrittämään tekijyyttä
- Selventää roolit ja vastuut sekä vähentää kiistoja
- On osa hyvää aineistojen hallintaa



**TUTKIMUSETIIKKA JA
AINEISTONHALLINTA -
TUTKIMUKSEN AIKANA**

MUISTA AINEISTOJEN KÄSITTELYSSÄ

- Käsittele aineistojasi aina eettisesti!
- Noudata lainsäädäntöä
- Muista käsittelyssä ja aineistojen siirrossa aina myös tietoturva
- Mikäli keräät ja käsittelet henkilötietoja, muista noudattaa EU:n yleistä tietosuoja-asetusta sekä yliopiston tietosuojapolitiikkaa
 - Kerää aineisto nimenomaista ja laillista tarkoitusta varten.
 - Älä kerää henkilötietoja, jos et tarvitse niitä tutkimuksessasi.
 - Minimoi henkilötietojen säilytysaika.
 - Suojaa tutkittavien yksityisyys anonymisoimalla data.
- Tampereen yliopiston tietosuojapolitiikka: <https://www.tuni.fi/fi/tutkimus/vastuullinen-tiede/hyva-tieteellinen-kaytanto>
- Katso myös tietosuojavaikuttetun ohjeistus tieteellisen tutkimuksen tietosuojasta: <https://tietosuoja.fi/tieteellinen-tutkimus>

AINEISTOJEN KUVAILU

- Tutkimusaineiston kuvailulla tarkoitetaan aineiston sisällön, keruun, muuttujien ja muiden tutkimuksen kannalta tärkeiden asioiden kuvaamista.
- Keskeistä on kuvailla tutkimusaineiston sisältöä ja rakennetta, ei niinkään tutkimuksesta syntynyttä julkaisua tai johtopäätöksiä
- Kuvailevaa tietoa voi tuottaa monella tapaa (tekstidokumentit, README-tiedostot, tekninen metadata, tms.)
- Kuvaile mieluummin liikaa kuin liian vähän
- Hyvin kuvailtu tutkimusaineisto on helpommin löydettävissä ja käytettävissä.
- Tavoite:
 - Ulkopuolinen ymmärtää miksi ja miten aineisto on kerätty, ja millä tavalla sitä voi käyttää omaan tutkimukseen
 - Aineiston tuottaja muistaa ja ymmärtää itse omaa aineistoaan

AINEISTOJEN SÄILYTYS

Arvioi seuraavia:

- millaista tutkimusaineistoa olet keräämässä ja tuottamassa?
 - Miten aiot käsitellä sitä? (tutkimusaineiston tyyppi ja määrä saattavat rajoittaa joidenkin säilytysvaihtoehtojen käyttämistä)
 - miten aiot tallentaa, säilyttää, käyttää, varmuuskopioida tai siirtää tutkimusaineistoasi?
 - kenen kanssa aiot jakaa tutkimusaineistosi?
 - miten tutkimusaineistoon pääsyä tarvitsee rajoittaa (pääsynhallinta)?
 - aiotko muokata tutkimusaineistoasi tutkimuksen aikana?
 - sisältääkö aineistosi sensitiivistä materiaalia? Sisältääkö se henkilötietoja?
-
- Suosi korkeakouluyhteisön palveluita aineistojen tallentamisessa ja säilyttämisessä
 - Muista koko datan käsittelyketjun tietoturvasuus!



**TUTKIMUSETIIKKA JA
AINEISTONHALLINTA
- TUTKIMUKSEN JÄLKEEN**

AINEISTOJEN AVAAMINEN

- Avoimen tieteen konteksti
- Tutkimuksen tausta-aineistot tunnustetaan merkittäviksi tutkimustuotoksiksi
- Avaaminen = aineistojen saattaminen muiden käytettäväksi ja hyödynnettäväksi
- Kun tutkimuksen tuotokset ovat laajasti hyödynnettävissä, tutkimuksen vaikuttavuus kasvaa
 - Tallennettu tutkimusaineisto meriitti cv:ssä
 - Kun data on tallennettu, myös muut voivat käyttää ja viitata dataan
 - Lisääntyneet viittaukset lisäävät vaikuttavuutta (Drachen et al. 2016; Piwowar, Vision 2013; Piwowar, Day & Fridsma 2007)
- Käytettävyyden lisäämiseksi olennaista on metadata eli kuvaileva tieto → ilman sitä jatkokäyttö hankalaa tai mahdotonta
 - Kuvailevaa tietoa voi julkaista mm. [Etsimessä](#)
 - Kansallinen Qvain-työkalu: <https://qvain.fairdata.fi/>
- Voit avata myös lähdekoodin esimerkiksi [GitHubissa](#) tai [Zenodossa](#)

AINEISTOIHIN VIITTAAMINEN

- Oikeaoppinen viittaaminen olennainen osa tutkimusetiikkaa ja osa hyviä tieteellisiä käytäntöjä
 - *”Tutkijat ottavat muiden tutkijoiden työn ja saavutukset asianmukaisella tavalla huomioon niin, että he kunnioittavat muiden tutkijoiden tekemää työtä ja viittaavat heidän julkaisuihinsa asianmukaisella tavalla ja antavat heidän saavutuksilleen niille kuuluvan arvon ja merkityksen omassa tutkimuksessaan ja sen tuloksia julkaistessaan.”* HTK-ohje
- Viittaa aina omaan tutkimusaineistoosi, mikäli se on avoimesti saatavilla
 - Aineiston jatkokäyttäjät voivat viitata aineistoosi
 - Lisää aineistosi näkyvyyttä ja löydettävyyttä
- Viittaa aina käyttämiisi aineistoihin!
 - Kunniaa aineiston alkuperäiselle tekijälle
 - Muut voivat etsiä saman aineiston käsiinsä
 - Hyödynnä arkistoivan tahon viittausohjeita

MISSÄ AVATA VASTUULLISESTI?

Käytä tieteenalallesi vakiintuneita paikkoja ja seuraa rahoittajasi ohjeita. Myös kustantajilla saattaa olla vaatimuksia julkaisuun liittyvän aineiston säilyttämisestä.

Tutkimusaineistoille soveltuva arkisto tai repositorio

- Valitse mahdollisuuksin mukaan luotettava arkisto (esim. Core Trust Seal –sertifioitu), esim. Tietoarkisto, Kielipankki
- Huomaa, on arkistoja ja ”arkistoja”

Datalehti

- julkaisevat esimerkiksi datankeruumenetelmiin ja datan laatuun liittyviä yksityiskohtaisia kuvauksia
- Springer Naturen Scientific Data
- Elsevierin Data in Brief
- Brillin Research Data Journal for the Humanities and Social Sciences

Vertaisarvioitu tieteellinen julkaisu

- Monet tieteelliset julkaisut edellyttävät tutkimusaineistojen avaamista tutkimustulosten verifiointiin ja läpinäkyvyyden lisäämisen vuoksi
- Tämän kaltainen aineisto ei kuitenkaan kaikissa tapauksissa ole avointa, jos siihen pääseminen edellyttää esimerkiksi lehden maksullista tilausta tai artikkelin ostamista
- Nykyisin monet kustantajat ohjaavat kuitenkin käyttämään datan säilyttämiseen tarkoitettuja paikkoja

FAIR DATA

FINDABLE – *löydettävissä*

Datalla on rikas metadata, joka sisältää pysyvän tunnisteiden

Data on indeksoitu hakupalveluun

ACCESSIBLE – *saavutettavissa*

Metadatan on oltava aina löydettävissä

Datan hakeminen avoimen, maksuttoman ja universaalin yhteiskäytännön kautta (mm. https-protokolla)

INTEROPERABLE – *yhteentoimivaa*

Datassa ja metadatatassa tulee käyttää hyväksytyjä formaatteja, kieliä ja sanastoja

Datan ja metadatan tulee sisältää linkkejä muihin datoihin ja metadatoihin

Järjestelmien yhteentoimivuus

RE-USABLE – *uudelleen / jatkokäytettävää*

Data on kuvailtu kattavasti metadatatalla

Datalla selkeät käyttöehdot

”REILUN” AINEISTON AVAAJAN MUISTILISTA

- Mieti jo tutkimusta suunnitellessasi, missä aiot julkaista datasi
- Käytä tiedostomuotoja, jotka ovat avoimia ja pitkäikäisiä
- Luo datalle metadata
- Varmista, että aineiston omistajuuteen ja tekijänoikeuteen liittyvät seikat eivät rajoita avaamista
- Määrittele aineiston käyttöoikeudet lisensoimalla aineisto
- Hanki datalle pysyvä tunniste
- Avaa aineisto sopivassa arkistossa tai datalehdessä
- Julkaise metadata

(How FAIR are your data? DOI:[10.5281/zenodo.1065990](https://doi.org/10.5281/zenodo.1065990))

- Muista **henkilötietoja** sisältävän datan kohdalla:
 - Voidaan avata vain anonymisoituna (pseudonymisoitu aineisto on edelleen henkilötietoa)
 - Käsittelyperuste voi rajoittaa jatkokäyttöä (esim. ei suostumusta)
 - Arkistoksi valitaan sopivan tietoturvatason arkisto
 - Vaikka aineistoa ei voisi julkaista, julkaise kuitenkin metadata
 - Hävitä data, jota et tarvitse



Tarjoamme yhden luukun periaatteella palveluita ja työvälineitä tutkimusdatan hallintaan Tampereen korkeakouluyhteisön henkilökunnalle ja opiskelijoille

Katso myös koulutukset: <https://research.tuni.fi/datapalvelu/koulutukset/>

LINKKEJÄ

Aineistohallinta

Research Data Management –verkko-opas: <https://libguides.tuni.fi/friendly.php?s=researchdatamanagement>

Tietoarkisto: Aineistohallinnan käsikirja: <https://www.fsd.tuni.fi/aineistohallinta/fi/>

Nerokkaan aineistohallintasuunnitelman tekeminen: <https://blogs.helsinki.fi/thinkopen/nerokkaan-aineistohallintasuunnitelman-tekeminen/>

Michener, William K. (2015): Ten simple rules for creating a good data management plan / PLOS Computational Biology: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004525>

Datan ja koodin jakaminen

Drachen, T.M. et al., (2016). Sharing data increases citations. *LIBER Quarterly*. 26(2), pp.67–82. DOI: <http://doi.org/10.18352/lq.10149>

Piowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE*, 2(3), e308. <http://doi.org/10.1371/journal.pone.0000308>

Piowar, H.A., & Vision, T.J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175; <http://dx.doi.org/10.7717/peerj.175>.

Vandewalle, P. (2012). Code sharing is associated with research impact in image processing. *Computing in Science and Engineering*, 14(4), 42–47. <http://dx.doi.org/10.1109/MCSE.2012.63>

Avoimuus

Tutkimuseettisen neuvottelukunnan HTK-ohje: http://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf

Ihmistieteiden eettinen ennakoarviointi (TAU): <https://www.tuni.fi/fi/tutkimus/vastuullinen-tiede/hyva-tieteellinen-kaytanto/ihmistieteiden-eettinen-ennakoarviointi>

EU:n avoimen tieteen suositus: <https://ec.europa.eu/digital-single-market/en/news/recommendation-access-and-preservation-scientific-information>

FAIR data: Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018, doi: 10.1038/sdata.2016.18, (2016).



OTA YHTEYTTÄ

Datapalvelu:

researchdata@tuni.fi

SEURAA

Twitter: @TampereUniLib

Facebook: Tampereen yliopiston kirjasto / Tampere University Library

Instagram: tampereunilib